Alexander, Robin (2012) Moral Panic, Miracle Cures and Educational Policy: what can we really learn from international comparison?, *Scottish Educational Review* 44 (1), 4-21

Moral Panic, Miracle Cures and Educational Policy: what can we really learn from international comparison?

The 2011 SERA Lecture, delivered at the annual conference of the Scottish Educational Research Association

Robin Alexander
University of Cambridge

ABSTRACT

In many countries, including the UK, the potential of international student achievement surveys such as TIMSS and PISA is being subverted by political and media fixation on the resulting league tables. These prompt not just well-founded efforts to learn from others' success but also ill-founded assertions about educational cause and effect, inappropriate transplanting of the policies to which success is attributed, and even the reconfiguring of entire national curricula to respond less to national culture, values and needs than to the dubious claims of 'international benchmarking' and 'world class' educational standards – the latter equated with test scores in a limited spectrum of human learning. Informing such responses are the attractively simple nostrums of high profile and highly selective literature reviews that massage policymakers' urge for the quick fix by playing down the complex interplay of culture and schooling and ignoring the kinds of evidence that can provide a truer and more nuanced picture of education systems in action. Using a typology developed by the US National Research Council, the paper critiques three recent and influential examples of this paradigm before illustrating an alternative approach. This draws on the author's comparative studies of culture and pedagogy to show how explicating the principles that underpin observed classroom practice, rather than copying national policies, can lead to genuine transformation of the quality and outcomes of student learning. The paper ends by contending that PISA panic and the supremacist mindset it feeds have dangerously distorted the debate about what a 'world class' education should entail. With PISA 2012 now in progress, policymakers are urged to redress the balance.

INTRODUCTION

At first sight, the title of this paper may not seem relevant to *Educational Research in an Age of Austerity,* the theme of the 2011 SERA Annual Conference. Actually it is, for the current 'age of austerity' is linked not only to the greed of maverick bankers but also to seismic changes in the global economy and to the growing economic might relative to Europe and North America of Brazil, India and above all China. With Shanghai, South Korea, Hong Kong and Singapore all now apparently untouchable at the top of the PISA league tables of student achievement in reading, mathematics and science, it's hardly surprising that Britain's policymakers believe that these Asian education systems have something to teach us. That is both sensible and responsible. It's *how* policymakers study education elsewhere and how they translate into policy what they discover that's the problem. Policy responses frequently display one or more of three tendencies: (i) extreme selectivity in the evidence cited; (ii) a misplaced faith in the ability of national interventions to transform classroom practice and hence raise standards; (iii) a preference for interventions that are high stakes and draconian – 'tough' is the usual epithet - presumably on the grounds that having publicly registered a problem policymakers believe that only the firmest of government firm hands will impress the electorate.

I make no apology for returning to a theme that has preoccupied me, on and off, since the mid-1990s (Alexander 1996, 2001, 2006, 2008a, 2008b, 2010a). For despite all that we

know about the pitfalls of cause/effect attribution in the educational and economic spheres, successive governments have found it hard to resist the naive belief that raising test scores in literacy and numeracy will elevate a country's economic performance, and that copying successful nations' educational policies will both raise standards and pull us out of recession. The prevailing term is 'leverage'. It's a revealing metaphor: depress one end of the lever (high-calory literacy and numeracy, high-octane school leadership, high-stakes tests) and up goes the other (improved educational standards and economic performance). If only it were that simple (Wolf 2002).

Thus the 'moral panic' of this paper's title is economic no less than educational, and the 'miracle cures' are the policies which politicians and school improvement experts nominate as the key ingredients in the success of the countries in question. Import these, they assert, and our problems will be solved.

Of course, what we are witnessing here in the millennial context of globalisation, economic turbulence and international re-alignment isn't new. At the end of the nineteenth century, as students of comparative education know well, Michael Sadler famously warned:

> In studying foreign systems of education we should not forget that the things outside the schools matter even more than the things inside the schools, and govern and interpret the things inside ... The practical value of studying in a right spirit and with scholarly accuracy the working of foreign systems of education is that it will result in our being better fitted to study and understand our own ... No other nation, by imitating a little bit of German organisation, can thus hope to achieve a true reproduction of the spirit of German institutions ... All good and true education is an expression of national life and character. (Sadler 1900, 50).

At the turn of the twentieth century all eyes were on German industrial might - with good reason as it turned out just 14 years later. Sadler's point was obvious and I'd have thought incontrovertible. National education systems are deeply embedded in national culture. Indeed education is one of the main vehicles for mediating, transmitting and sustaining a culture. So, as I've argued elsewhere, 'no educational policy or practice can be properly understood except by reference to the web of inherited ideas and values, habits and customs, institutions and world views, that make one country distinct from another.' (Alexander 2001, 5).

This means that while, in Sadler's words, 'All *good and true* education is an expression of national life and character' (my italics – see below), it is also the case that, in the words of American educationist and adminstrator Ernest Boyer (1983, 6), 'A report card on public education is a report card on the nation. Schools can rise no higher than the communities that support them.' Or, as Basil Bernstein more aphoristically expressed the matter 40 years ago: 'Education cannot compensate for society.' (Bernstein 1970, 344)

That is to say, if a national education system is not 'good and true' then we must seek explanations in our society as well as our schools. Thus it is, as is well known but insufficiently acknowledged, that the gap in student achievement between high and low-attainers is wider in Britain than many other rich nations in part because it maps with depressing precision onto the gaps in income, health and wellbeing which are also wider in Britain than in most other rich nations. As Wilkinson and Pickett (2010) demonstrate in their groundbreaking epidemiological study *The Sprit Level*, unequal societies have unequal education systems and unequal educational outcomes. Britain and the United States illustrate these interlocking inequalities *par excellence*, though in this context 'excellence' is hardly the right word.

Obvious all this may be, yet it's not obvious enough to the current crop of educational policy advisers who advocate policy borrowing on what I believe is a culturally reckless scale and are applauded by their governments for doing so; and who find it politically more expedient to counter Britain's patchy educational performance by commending the educational policies of Singapore, Hong Kong or Finland than by examining the economic and social policies of Britain.

To be fair, though historical awareness is something that such discourse conspicuously lacks, history also sets limits to the purist Sadler line quoted above:

> Cultural borrowing happens; it has always happened. Few countries remain hermetically sealed in the development of their educational systems, and for centuries there has been a lively international traffic in educational ideas and practices. So, for example, Pestalozzi mingles with Tagore, Krishnamurti and the Elmhirsts in both English and Indian progressivism; Dewey turns up briefly in China, the Soviet Union and Turkey as well more lastingly in England and the United States; both the German *Gymnasium* and the American high school help shape the development of Russian schooling; Kay Shuttleworth imports or exports the *Ecole Normale* from France to England and India; Jan Amos Komensky (Comenius) journeys tirelessly from Moravia to Prague, Heidelberg, Amsterdam, Berlin, Paris, Stockholm, London and points between and beyond, and corresponds with the founding fathers of Massachusetts, and his principles of common vernacular schooling and carefully graduated whole class teaching, not to mention his textbooks, embed themselves deeply and lastingly in the pedagogy of many countries of central, eastern and northern Europe; and the monitorial systems of Bell and Lancaster seed themselves just about everywhere from their probable roots in what was then Madras. (Alexander 2001, 171-2).

The question begged by this apparent paradox is this: if culture is central to education, and cultural transplants are on the face of it doomed to failure or at best limited success, what exactly is the character of that educational thinking and practice which, over the past thousand years or so, has migrated *successfully* from one country to another and embedded itself there? By the end of this paper I hope to have answered this, in part by reference to my own efforts to apply the lessons of international comparison to the improvement of classroom practice in England. But we might also note the real force of Sadler's claim that 'the practical value of studying in a right spirit and with scholarly accuracy the working of foreign systems of education is that *it will result in our being better fitted to study and understand our own*.' (My italics). The prime purpose of comparative international study, Sadler insists, is to understand ourselves, not to copy others. Or as David Raffe (2011a) argues, 'policy learning' is a more valid and effective pursuit than 'policy borrowing'.

In this paper I shall examine the current educational manifestations of moral panic and miracle cures by considering four questions:

- What kinds of evidence are policymakers most interested in?
- What do they do with the findings?
- What is wrong with what they do?
- How might policymakers make better use of international comparison?

WHAT KINDS OF EVIDENCE ARE POLICY MAKERS MOST INTERESTED IN?

The official answer to this question is 'evidence which will help them develop worthwhile and effective policies.' The cynical answer is 'evidence which is carefully selected to justify policies that have already been determined.' This invokes the familiar jibe that what is presented as 'evidence-based policy' is likely to be policy-based evidence. The jibe may sound cheap, but it has substance - as is shown by the evidence and experience of the Cambridge Primary Review (Alexander 2010b, Alexander *et al* 2010, Alexander 2011).

But beyond policymakers' preference for evidence that supports chosen policies rather than questions them, there is also a tendency to pay greater attention to some research paradigms and forms of data than to others. Take, for example, data from international comparison. Among the many typologies of comparative international research – descriptive/prescriptive, quantitative/qualitative, single country/two or more countries, macro/micro/macro-micro and so on – one of the most useful in the context of its policy applications has come from the National Research Council (NRC) of the US National Academies. Through its Board on International Comparative Studies in Education (BICSE - sadly no longer extant, though its work continues), the NRC has both advised and

commented on US participation in the international student achievement surveys over many years. The 2003 NRC/BICSE report differentiates three main types of comparative study:

> Type I studies typically include large-scale surveys that aim to compare educational outcomes at various levels ... Type II studies are designed to inform one or more particular ... education policies by studying specific topics relevant to those policies and their implementation in other countries. Type III studies are not designed to make direct comparisons ... in terms of specific policies or educational outcomes. Rather, they aim to further understanding of educational processes in different cultural and national contexts. (NRC 2003, 13)

Thus Type I includes the large-scale international student achievement studies like TIMSS, PISA and PIRLS. These are typically quantitative and are referenced by a limited number of indicators and measures of learning outcome and, latterly, of context and input.

Type II includes a more varied mix of studies though the usual paradigm is the desk-based review of existing literature or data. What Type II studies have in common is a direct focus on specific policy concerns – standards, the curriculum, teacher training, school leadership or whatever. Examples are the recent reports from Ofsted, McKinsey and Cambridge Assessment which adopt the Type II desk review format first to explain countries' differential performance in Type I student achievement surveys and then to propose policy responses. (Reynolds and Farrell 1996, Barber and Mourshed 2007, Mourshed, Chijioke and Barber 2010, Barber, Whelan and Clark 2011, Oates 2010). Type II desk reviews, it seems to me, most conspicuously illustrate the dangers of ill-conceived international comparison in the context of educational policy, so I shall use them as cases.

Type III includes the majority of work in the published corpus of academic comparative education. These studies range from pedestrian accounts of the education system of countries x or y to truly illuminating comparative studies of classroom life and what shapes it. Good examples of the latter are Joseph Tobin's *Preschool in Three Cultures* and *Preschool in Three Cultures Revisited*, and the Bristol group's *World of Difference?* exploration of the experience of being a learner in three European countries. (Tobin *et al* 1989, 2009, Osborn *et al* 2003). Tobin's 'in three cultures' signals the much greater effort entailed here than in the bog-standard account of educational system x or y produced by the academic hacks of comparative education. Type III territory is uneven, and the scale, range and methods of its instances are variable, but what all such studies have in common is that though some may have significant policy applications the imperatives of policy are not their principal impetus. Rather, their goal is the advancement of understanding of other countries, their education systems, schools and classrooms, for its own sake. But even when Type III studies do have policy applications they tend to be ignored by government because, in that phrase redolent of ministerial complacency and civil servant defensiveness, they are 'not invented here'. Which is to say, if they are not commissioned or approved by government, or if they cannot comfortably be plagiarised, then they are assumed to have nothing to offer.

Thus it is that governments endlessly confirm and recycle their own worldviews and policies, and the follies that attend them. Confirming this, the 2003 NRC report notes that while the majority of published comparative education studies are Type III, it is the Type I and II studies that receive most of the funding, political patronage and/or publicity, and the funding difference per study can be truly vast. Type I studies in particular are a multi-million dollar business, Type II may be generously or modestly funded, depending on whether they entail fieldwork or merely a literature review. Unless they come from McKinsey, that is. Tim Oates' Cambridge Assessment paper 'Could do Better?' (Oates 2010) which has so impressed England's education ministers, is a standard black-and-white laser printer job. But the first McKinsey report (Barber and Mourshed 2007) is a lavish coffee table product with colour illustrations, an exotic typeface, and cardboard covers of the robust grade normally reserved for eco-coffins. It is also so big that you have to stand up to read it – an act of enforced deference that I somewhat resent. And all the McKinsey education reports have been backed by a no less prodigal campaign of international promotion. In stark and pitiful contrast, Type III studies scrape together what they can from charities and research councils

and are disseminated through decidedly low-profile educational journals. Yet, the NRC report goes on:

> Although they vastly outnumber Type I and Type II studies, Type III studies do not usually come to the attention of policy makers or the public. This is a loss, since many are rich in narrative detail and paint a more engaging and provocative portrait of education in other countries than do the summary bar charts and graphs typical of many larger studies. Ethnographic and case studies, in particular, can explore cultural context in depth and, in turn, help elucidate the way education is organised and understood in different cultures. (NRC 2003, 23-4).

Yet it must also be acknowledged that the failure of policymakers to attend to Type 3 studies is partly the fault of the academic community itself. Such studies may lie buried in low-circulation, closed-access and overpriced academic journals, and too few education academics accept the need to communicate their work in accessible form to policymakers and practitioners. Many, indeed, prefer to remain within their institutional cocoon and are fearful of sticking their heads above the parapet.

WHAT DO POLICYMAKERS DO WITH THE EVIDENCE?

Having funded and studied Type I studies like TIMSS and PISA, and having commissioned or welcomed Type II literature reviews that seek to explain the resulting league tables and propose what to do about them, and having ignored the Type III studies which may offer the insight they desperately need but would prefer not to be inconvenienced by, what do our policymakers do next? I answer that question by referring to the three high-profile examples I mentioned earlier.

In 1996 Ofsted, England's equivalent of HMIE, published a report commissioned from David Reynolds and Sean Farrell entitled *Worlds Apart? A review of international surveys of educational achievement involving England.* This reviewed England's performance in the international student achievement surveys published up to that point. It then proposed reasons for England's generally poor performance, and identified solutions. Its principal conclusion was that the complex pedagogy of English primary classrooms, with its emphasis on grouping and individualisation, accentuated the already wide differences between pupils. To counter this the authors recommended:

> High quantities of whole-class interactive instruction, in which the teacher attempts to ensure the entire class have grasped the information being given … the use of the same textbooks by all children … mechanisms to ensure that the range of achievement is kept small. (Reynolds and Farrell 2006, 55).

The judgements about pedagogical complexity and children growing apart were pertinent and indeed had been anticipated in a number of earlier Type III studies (Galton and Simon 1980, Mortimore *et al* 1988, Alexander 1997), though the real political appeal of *Worlds Apart?* to right-leaning governments, whether Conservative or New Labour, was its apparent invitation to go back to basics with didactic whole class teaching and standardised textbooks.

Next, in 2007, the multi-billion dollar international management consultancy McKinsey published the first of three reports from a team headed by Michael Barber, formerly of Prime Minister Blair's back office, on what the education systems that performed best in PISA could teach those which by this single debatable criterion are judged less successful. Its provocative title was *How the World's Best-Performing Education Systems Come Out on Top*. From the ten top performers in PISA 2003 it concluded:

> Three things matter most: 1) getting the right people to become teachers, 2) developing them into effective instructors, and 3) ensuring that the system is able to deliver the best possible instruction for every child. (Barber and Mourshed 2007, 2).

In the course of its expensive quest for this damp squib the McKinsey report delivered further stunning insights such as: 'The quality of an education system cannot exceed the quality of its teachers', 'The only way to improve outcomes is to improve instruction' and 'High performance requires every child to succeed.' To such compound tautologies McKinsey added linguistic gems like: 'Top-performing systems leverage a substantial and growing knowledge about what constitutes effective school leadership to develop their principals into drivers of instruction.' (Barber and Mourshed 2007, 4 and 30).

Two further reports followed from the same stable and in similar methodological and linguistic vein: *How the World's Most-Improved School Systems Keep Getting Better* (Mourshed *et al* 2010) and *Capturing the Leadership Premium: how the world's top school systems are building leadership capacity for the future* (Barber *et al* 2011). The second McKinsey report looked at 20 rapidly-improving systems, the third at eight. 'Systems' in all three reports was defined very loosely indeed, allowing comparisons, which some would balk at, between Japan (a country with a population of 127 million), Alberta (a Canadian province with 3.7 million) and Aspire (a charter school system in the state of California with just 40 schools).

The third case is the report by Tim Oates, of the examination agency Cambridge Assessment, which has a foreword by England's Secretary of State for Education, no less, and is entitled *Could do Better: using international comparisons to refine the national curriculum in England* (Oates 2010). Notice the preference in these Type II studies for titles which have the appropriate back to basics ring of old-fashioned school reports – 'worlds apart', 'how to come out on top', 'could do better', 'most improved', though what these documents more immediately bring to mind is the most entertaining if not significant educational report of my childhood, by one Nigel Molesworth: *How to Be Topp: a guide to sukcess for tiny pupils, including all there is to kno about space*. Sadly there are no Ronald Searle illustrations to enliven the other reports I've mentioned.

But I digress. With Oates as with the others, the most recent international student achievement survey (in this case PISA 2009) provides the 'benchmark' and on the basis of a curious range of sources, with not a single significant Type III study in sight, Oates concludes:

> In all high-performing systems, the fundamentals of subjects are strongly emphasised, have substantial time allocation and are the focus of considerable attention. (Oates 2010, 10).

Oates goes on to list 13 'control factors for transnational analysis'. His list is so long and diffuse that it confronts governments with the choice of ignoring it or adopting it in full, which would require them to take over every school in the land. Conspicuous by its absence from the list of 'control factors' (ominous phrase) is culture: more on this anon.

Thus we have three reports, all enjoying classic Type II political patronage (in McKinsey's case by many more governments than that of the UK), all starting from and giving unquestioning credence to the same kind of dataset – international student achievement surveys covering a very limited range of outcomes in relation to the totality of what most schools seek to achieve – yet, interestingly, each reaching different conclusions. Looking at high-performing systems as these are defined by their showing on the international league tables of student achievement, educational improvement requires, variously: more whole class teaching coupled with standardised textbooks (Reynolds), better teaching, teacher training and school leadership (McKinsey), and a drastically pared-down curriculum which concentrates on what is deemed 'essential knowledge in the key subject disciplines' (Oates as cited in the remit for England's 2010-12 national curriculum review, DfE 2010a). Whole class teaching (Reynolds) and a narrow subject-based curriculum (Oates) fall into my category of 'miracle cures'. The McKinsey recommendation of better teachers, teacher training and school leadership is decidedly unmiraculous since it is a statement of the obvious.

What have policymakers done with these three studies? The 1996 Ofsted study's advocacy of whole class teaching and standardised textbooks contributed directly to the pedagogy of the national literacy and numeracy strategies introduced by the Labour government in 1998-9 which for a decade were imposed on every primary teacher in England and policed by Ofsted, only to be abandoned by the coalition government in 2010. The Oates paper led to Oates's appointment as leader of the so-called 'expert panel' which is currently advising the UK government on how to slim down England's national curriculum to what is 'essential'. In December 2011 this group produced their report (DfE 2011a). Notwithstanding the report's attempt at educational rationale, the slimming-down exercise is chiefly one of placing more subjects below a line notionally marked 'essential' than were below that line in the previous version of England's national curriculum. Crucially, 'essential' is not defined, except by reference to what 'high performing jurisdictions' do.

The impact of McKinsey is more diffuse. It can be seen, for example, in the UK government's raising of the academic entry bar for trainee teachers (in England), though nowhere near as high as Finland, the country which is cited in this context. However, while saying that good teaching is what makes the difference, McKinsey had nothing whatever to say about pedagogy as such, so its reports have no obvious *classroom* purchase, and in any case the focus is on systems rather than classrooms. But the McKinsey brand is evident in the current elevation of school leaders from head teachers, who by their traditionally modest English title are *primus inter pares*, to heroic figures who turn round failing schools and stamp their sharp-suited presence on the mere mortals who actually do the teaching, and who - as John Bangs of the England's National Union of Teachers has observed - seem in danger of being forgotten altogether (Bangs 2011). Instead, what appears to be emerging as sanctioned policy is a staffing structure for schools – especially large secondaries – which is decidedly top-heavy. Stephanie Northen has charted this trend, which in one 1,300-pupil secondary school has yielded a principal, a senior vice-principal, two vice-principals, eight assistant principals and a finance director, all commanding the kinds of salaries which hitherto have been confined to the business sector - whose style, clearly, all this seeks to replicate. (Northen 2011).

Incidentally, this decidedly Anglo-Saxon view of school leadership has become another post-PISA miracle cure, recommended by McKinsey for universal consumption despite the fact that in many countries – including top-performing Finland - school heads/principals/directors have a much more modest role than in Britain and are by no means free to impose their will on other teachers.

WHAT IS WRONG WITH WHAT POLICYMAKERS DO WITH INTERNATIONAL EVIDENCE?

If Type II policy extrapolations could be shown to have raised standards there would be no grounds for complaint. So far, however, it's not at all clear that this is so, and the reasons for this lies partly in the methods the authors use and the inferences they draw.

As to the Type I datasets which provide Type II studies with their benchmarks, from the methodologically shaky IEA and IAEP studies of the 1970s to the much more sophisticated 2009 PISA, these have become much more reliable and culture-fair, and the most recent PISA commentaries produced by OECD demonstrate considerable awareness of the challenges of cross-cultural testing and the inferential and explanatory caveats that need to be entered and heeded. (OECD 2010a, 2010b, 2010c, 2010d, 2011a, 2011b).

In this sense, the problem isn't so much PISA as what, faced by the resulting league tables and the surrounding media noise, policymakers and their advisers do with what PISA provides. Hence my particular concern about the Type II data extrapolations and the disproportionate influence they exert. For the political attraction of Type II studies is that they select, mediate, repackage and re-interpret the research of others, presenting it in a form which they believe policymakers will find palatable both politically and stylistically. Being gatekeepers rather than creators of evidence, Type II studies are acutely vulnerable to the charge of methodological and/or ideological bias.

There are now available some pertinent critiques of the Ofsted and McKinsey studies though as yet there's disappointingly little on the Oates paper. Although it seems to me that the 1996 Ofsted *Worlds Apart?* report (Reynolds and Farrell 1996) is by far the best of the bunch, both because it has genuine grounding in what actually happens in schools and classrooms and because it sounds appropriate notes of caution to which the McKinsey reports are immune, I remain critical of this one too. In the Ofsted study, for example, the quality and effectiveness of whole schools and entire education systems is reduced to a statistical calculation of gain in output over input. The chosen measures of input and output are extremely restricted in relation to what we know from other sources about the contexts, conditions, processes and outcomes of schooling and learning. Output measures are confined to students' test scores in limited aspects of a narrow range of subjects, and these are taken as proxies for pupil attainment across the entire curriculum. The 'process' measures which are added to the mix in order to calculate what aspects of education make a difference are no less restricted, for they must satisfy the basic requirement of measurability - hence the fixation on measures like time on task (which the late Nate Gage called 'a psychologically empty concept'). (Gage 1978, 75). Most egregiously, culture - which is absolutely central to the proper pursuit of educational comparison - is reduced to one 'factor' among many, something which is external to school life rather than that which creates it and gives it meaning. Finally, the literature on which the paradigm draws represents a very narrow segment of the wider literatures on comparative and international education and on school and classroom processes. (Alexander 2001, 29-30 and 36-9).

All these criticisms apply *a fortiori* to the Oates and McKinsey Type II reports. In my 2009 presidential address to the British Association of International and Comparative Education (BAICE), I argued that McKinsey falls well short of the kind of analysis which is required by a study purporting to explain how PISA high-flyers achieve their success. (Alexander 2010a). For example, McKinsey (Barber and Mourshed 2007) says that the quality of teaching matters most, and we'd probably agree that it is critical though the balance of within-classroom and other factors is very much open to debate. Yet, almost perversely it seems, they offer no account or analysis of teaching whatever, conceptually or empirically. Further, like the 1996 Ofsted study, McKinsey's methodological repertoire and vocabulary start and end with the input-output preoccupations of school effectiveness research, and the authors display unshakeable faith in the power and validity of testing. 'All of the top-performing systems', say McKinsey, 'recognise that they cannot improve what they do not measure' (Barber and Mourshed 2007, 36), which presumably means that the curriculum beyond English, maths and science is of no account. Beyond that, McKinsey relies almost totally on official information about its chosen education systems, which can be pretty unreliable because the ministries that supply it have an interest in presenting their systems and policies in the best possible light. Put these objections together and you have a study which is almost bound to be weak, conceptually and methodologically. Add to the mix the report's abundance of cliché and plain bad English, and you have a document that falls at the early hurdles of veracity, validity and meaningfulness.

Since, in spite of or possibly because of these failings, McKinsey's three education reports are currently proving so influential internationally, we might press the critique further. David Raffe (2011b) argues that the second McKinsey report remains locked within the policy 'bubble'; that it is a particularly flawed example of policy borrowing; that it makes light of all-important educational processes and contexts; and that it applies exceptionally narrow criteria of systemic educational success. But the most devastating criticisms come in a recent article by Frank Coffield (2012). His central charge is that McKinsey's analysis is culpably mono-factorial when it has long been understood that socio-economic factors have a significant impact on students' motivation, engagement, learning and attainment, especially in a country whose economic and social disparities are as great as they are in Britain. The first McKinsey report, continues Coffield, is methodologically flawed, 'disablingly selective' in its data and explanatory frame, superficial in its account of 'best practice' and how this can be disseminated, and seduced by its own rhetoric on leadership. Coffield likes the second McKinsey report a bit better than the first, though still not a lot, but castigates it

for an impoverished view of teaching and learning, a thin evidence base, implausible arguments about the mechanisms and processes of school improvement, technocratic and authoritarian language and a pervasive neglect of culture and political context. Yet, Coffield notes, the UK coalition government's 2010 White Paper *The Importance of Teaching* (DfE 2010b) approvingly quotes McKinsey seven times in its first 20 pages. Coffield sees McKinsey as

> the work of 'global' policy analysts, remote from the complexities of classrooms and the discomfiting findings of researchers which pose such difficulties for politicians in search of quick 'transformations' of school systems before the next election. They espouse a ... model of schooling ... characterised by relentless pressure, competition, line managers, customer services, data for performance management, accountability and value for money; and professional autonomy for teachers only when granted by the centre ... Their notion of teaching is narrowly conceived and technocratic ... Their model remains unsophisticated, impracticable and undemocratic ... Their recommendations are educationally and socially dysfunctional and should not be part of school reform in a democracy. (Coffield 2012, 145-6).

Strong stuff indeed, though in my view not unjustified. The paradox is that McKinsey's authoritarian and micro-managerial nostrums are welcomed by the same government that says it wants to give teaching back to teachers after 13 years of what in the Cambridge Primary Review final report we called a 'state theory of learning.' (Alexander 2010b, 291).

To the criticisms of Raffe and Coffield I'd add others which seem to me to be no less fundamental.

Type II comparative studies of the kind I have exemplified are desk-based and therefore lack the vital asset of first-hand empirical data systematically and transparently presented, and analysed in accordance with methodological procedures that are publicly recognised or, if they are experimental, are at least open to scrutiny. These three studies are notable for a high degree of selectivity, arbitrariness and bias in the literature on which they draw and in the way they handle it. That, I'm afraid, may be part of their attraction in policy circles, because they select their evidence to prove a point or come up with a politically-acceptable solution rather than demonstrate the complexity which a properly-conducted comparative analysis is likely to reveal.

Aside from the results of PISA 2003 and 2009, the core database for these studies is information that has been generated by governments or their agencies. From such material - high on political rhetoric, low on classroom practice, devoid of cultural context, cleansed of problematic realities – they then construct edifices for school improvement and systemic reform. Disarmingly straightforward and carefully pitched to appeal to political instincts, these studies are eagerly taken up in the corridors of power.

The ready acceptance of official statements and the neglect of school and classroom practice are all the more serious when we note from Hogan's work in Singapore that in a system which outperforms most others in PISA - and whose policies both Barber and Oates urge governments elsewhere to copy - there is a substantial gap between the prescribed and the enacted curriculum despite the existence of mechanisms to secure coherence and control, and the sources of this gap are 'multiple, mutually reinforcing, resilient and intractable'. Hogan adds that using system-level curriculum specifications and controls as levers for reform, as Oates recommends, ignores what really makes the difference. (Hogan *et al* 2012). Yet Oates's entire thesis centres on the curriculum as officially specified, and its hegemony and presumed efficacy are such that this considerable limitation isn't even mentioned. Similarly, one of the key supporting documents for England's national curriculum review is subtitled *What can we learn from the English, mathematics and science curricula of high-performing jurisdictions?* and consists of an exhaustive analysis of official syllabuses in nine systems. (DfE 2011b). For policymakers, 'curriculum' is what they prescribe, not what teachers do.

And what of the measures that prompt policymakers to summon the dispensers of miracle cures? PISA assesses the attainment of 15 year olds in aspects of reading,

mathematics and science. Its spectrum of 'key competencies', though clearly essential, is actually quite limited. But to its credit, and unlike some who use it as a badge of national superiority or a stick with which to beat teachers, PISA itself more realistically and modestly acknowledges that its tests cover just 'some of the knowledge and skills that are essential for full participation in society'. (OECD 2012). PISA is right, and Reynolds, Barber, Oates and gullible government ministers worldwide should pay attention: the TIMMS and PISA surveys are not a sufficient basis for describing an education system as a whole as 'high performing'.

Next, we return to what some call 'cherry picking', a tendency that publicly all deplore but many continue to pursue. This has a statistical variant: false correlation, or a version of the philosophers' 'fallacy of division'. X may well be a common feature of high-performing education systems a, b, c, d and e, but that doesn't demonstrate a cause-effect relationship between feature and performance. And if x is also a common feature of *low*-performing systems g, h, i, j and k, then the claimed relationship is clearly inadmissible. In 1996, the Ofsted *Worlds Apart* study (Reynolds and Farrell 1996) found that in its chosen high performing systems (at that stage as judged by TIMSS results) whole class teaching was the main teaching method used. Reynolds and Farrell therefore concluded that whole class teaching was one of the keys to delivering high standards, hence its commendation in the Labour goverment's national literacy and numeracy strategies introduced in 1998 and 1999. However, had they looked at other systems they would have discovered that whole class teaching is the international pedagogical default, as prevalent in low-performing systems as high, so actually there is no correlation.

The same applies to Oates's claim that the key ingredient of success in PISA is a curriculum concentrating on what is deemed to be 'essential knowledge in key subjects'. In fact, like whole class teaching and as the Benavot studies and International Review of Curriculum and Assessment Frameworks (INCA) curriculum databases show, a curriculum constructed in terms of 'essential knowledge in the key subjects' is the basis of most of the world's national curriculum specifications, PISA successes and failures alike, though of course they may differ in their view of which subjects are 'key' and what knowledge is 'essential'. (Benavot *et al* 1991, Benavot 2008). In short, far from being unique to high-performing systems, Westminster's current curriculum holy grail of a narrow range of subjects dominated by literacy, numeracy and science is a global curriculum commonplace. As with whole class teaching, there is no correlation between this account of curriculum structure and student achievement scores in TIMSS, PISA or PIRLS. The issue, of course, is what schools and teachers *do* with national curriculum specifications, which is about pedagogy, the area which the Oates and McKinsey studies ignore, much more than curriculum. Hogan's recent findings on the prescribed/enacted curriculum in Singapore (Hogan *et al* 2012) confirm this. For Oates, pedagogy is relegated to the status of one 'control factor' among many, acquiring the same lowly and almost incidental status as culture in the Reynolds and Farrell study.

Once we move beyond the restricted range of system and school variables deployed in these Type II studies, we encounter wider social, cultural, demographic and economic conditions which directly and massively influence the educational performance of a country's students. In a recent critique of such studies I show that while we can rightly conclude that teachers and teaching make a considerable difference (true but banal and probably not worth McKinsey-level expenditure and hype), extra-educational factors like country size, per capita GDP, demographic homogeneity and relative equality may well correlate no less convincingly with PISA performance. Indeed, a glance at the top end of the PISA league tables shows that the systems in question are mostly small, rich or preferably both, and to these factors we can add another from Wilkinson's and Pickett's groundbreaking epidemiological study *The Spirit Level*:

> Greater equality, as well as improving the wellbeing of the whole population, is also the key to national standards of achievement and how countries perform in lots of different fields ... There is not one policy for reducing inequality in health or the educational performance of school children, and another for raising national standards of performance ... If ... a country

> wants higher average levels of educational achievement among its school children, it must address the underlying inequality which creates a steeper social gradient in educational achievement. (Wilkinson and Pickett 2010, 29-30)

Interestingly, in their most recent survey PISA itself confirms the significance of equity as a factor in national performance. (OECD 2010d, 9-11).

So there's a constellation of factors in which wealth, demography, equity and relative equality all play a part alongside the particular school and education system factors on which our three Type II studies concentrate, though in the end it's culture which determines how wealth is disposed, how education is conceived and how much or little equality matters. On all such matters, the Ofsted, McKinsey and Cambridge Assessment reports are strangely silent.

HOW MIGHT POLICYMAKERS MAKE BETTER USE OF INTERNATIONAL COMPARISON?

The immediate but perhaps uncharitable answer is that policymakers should use researchers who are less prone to the basic errors I have exemplified and who are more careful in the inferences they draw, the explanations they offer and the policy options they recommend, and who tell policymakers what they need rather than what they wish to hear. Research in heavily contested policy arenas such as education must always and unflinchingly strive to speak truth to power.

The second thing to say is that once one entertains a wider range of explanatory factors in international student achievement than the limited spectrum considered by the three influential studies of educational cause and effect that I have discussed, the possibilities for genuinely useful international comparison increase dramatically.

For example, in 2007 UNICEF published a study placing the UK at the bottom of a league table of children's wellbeing in 21 rich nations, using indicators for income, health and safety, family and peers, behaviour and risk, and of course education, plus one for the subjective judgement of children themselves. (UNICEF 2007). A follow-up study explored the reasons for Britain's poor showing in the 2007 study by comparing the UK, Spain and Sweden. (UNICEF 2011). It concluded that a materialistic and commercial culture is deeply embedded in the UK and in concepts of good parenting in a way that is not seen in Spain and Sweden, and this culture reinforces inequality, adversely affects family time and relationships, and overall has a negative impact on children's wellbeing. Since we know that the maps of attainment and social equity broadly coincide, we have in such studies clues to how schooling, through the values and relationships it fosters as well as through its quality of teaching, might raise the standards of both attainment and wellbeing.

Here's another example. My Cambridge colleague John Gray has recently completed a meta-analysis of adolescent wellbeing, using international data, which shows that while the Netherlands and Finland both rank high on PISA, Dutch students are much happier at school than their Finnish contemporaries. He also argues that McKinsey-style comparisons between the UK and Hong Kong or Singapore are fruitless because the cultures are so utterly different and the educational systems are not remotely comparable in scale (England has 23,000 schools while Singapore has just 350). He therefore suggests that if we really wish to use international comparison for policy transfer we should look not at Hong Kong, Singapore or even Finland, but at the Netherlands, because it is successful in the interlocking areas of attainment and wellbeing and it is culturally not too different from the UK. (Gray *et al* 2011).

Another Cambridge colleague takes a different but no less intriguing view. Like myself, David Hargreaves sees culture as central to the understanding of differences in educational structure, content, process and outcome, but doesn't follow Gray's line on how we should act on that understanding. For Hargreaves, the world's educational and economic centre of gravity in the 21st century will be China, and since Shanghai-China outperforms most other countries in PISA it is essential that we investigate what it is about Chinese culture that

accounts for this and what we can extrapolate not so much from the education system as from the culture that shapes it. This inevitably leads him to the resilience and pervasiveness of Confucianism, whose values he suggests might profitably inform our own schools' collective identity and their approach to self-improvement. (Hargreaves 2011).

Though Confucius invariably features in explanatory accounts of East Asian approaches to education (Watkins and Biggs 1996, Tobin *et al* 1989 and 2009, Li, 2012), this is a novel and on the face of it startling response to Sadler's maxim that 'all good and true education is an expression of national life and character': import the culture, not the policies. On the other hand, Hargreaves makes a useful distinction between policy *replication*, *adaptation, grafting* and *redesign* and suggests that only the last of these has much chance of success, though it is also considerably more difficult to achieve than the other three. (Hargeaves 2011, 6-7). This is closer to my own argument that policy learning from international comparison should be about explicating and where appropriate emulating *values and principles* rather than importing and copying *policies or practices*. But this requires a comparative methodology capable of distinguishing the lived-in values and principles of education in action from the erzatz or cosmetic claims of policy and official documents, and this points us decisively away from study Types I and II, and as decisively towards the aspirations, if not always the achievements, of Type III.

FROM BELGOROD TO BARKING AND KURSK TO KIRKBY OVERBLOW: PRINCIPLED POLICY LEARNING AT WORK

I have an obligation to take up the challenge I have thrown down. In 2001 I published *Culture and Pedagogy*, a large-scale Type III comparative study of the relationship between history, culture, policy, schooling and classroom practice in primary education in England, France, India, Russia and the United States (Alexander 2001). The study culminated in close-grained analysis of interaction and discourse in the classrooms of these countries, drawing on a dataset that included 130 hours of videotape, transcripts of 166 lessons in over 100 schools, together with fieldnotes, photographs, interview transcripts, official documents, teachers' lesson plans, students' written work, and much else. The discourse analysis was the culmination of the study both in the literal sense that this was where it ended, but more fundamentally because it is through language, and especially spoken language, that culture, teaching and learning are most decisively yet subtly mediated.

When at the end of the study I asked 'What can English primary education, and English pedagogy, learn from the pedagogy of France, India, Russia and the United States?' (Alexander 2001, 563), I sounded the usual cautions about policy borrowing but then argued that precisely because spoken language is so central to both human learning and collective culture and identity, and precisely because the differences I had observed were so striking, classroom talk surely offers a rich potential for policy learning, one far richer than the systemic structures and policies commended by McKinsey, let alone the dozens of national curriculum prescriptions laboriously collated by the 2010-12 DfE national curriculum review. For in the end - I say yet again - it's pedagogy that makes the difference. Accordingly I embarked on the task which over the past decade has led to what I call 'dialogic teaching' being trialled and adopted in a large number of schools and teacher education courses, mainly in Britain but in other countries too.

Dialogic teaching (Alexander 2008a, 2008c) is part of the wider movement of classroom talk reform whose notables include Courtney Cazden, Lauren Resnick, Martin Nystrand, Sarah Michaels and Cathy Connor in the United States, and, in Britain, Douglas Barnes, Tony Edwards, Neil Mercer, Phil Scott and Philip Adey. But whereas some of these have tended to be impelled by reaction against the continuing dominance of the initiation-response-evaluation (IRE) 'recitation script' of closed questions, recall answers and judgemental rather than informative feedback, I have been able to add cross-cultural perspectives and pedagogical principles and strategies which at first glance may seem alien but once explored and applied work rather well.

Thus from Russian and French classrooms I was able to identify and then challenge the still pervasive Anglo-American belief - manifested in both teachers' accounts of their practice and videorecordings and transcripts of their classroom discourse - that the main function of classroom talk is *social* rather than *cognitive*, or that it's about boosting the child's confidence rather than getting him or her to think. This view of talk conditions the organisational principle that in a given lesson every child must be persuaded to say something. Once one calculates the logistics of this strategy in relation to class size and lesson time, it's evident that it is more likely to secure participation in a lesson's social round than in the learning the lesson seeks to advance, because such participation is likely to be restricted to brief exchanges that have shallow cognitive content or impact (and indeed the data show this to be true - Alexander 2001, pp 391-528). In the Russian and to a lesser degree the French classrooms - and since then I've observed this approach in several other countries - teachers consciously opted for extended and more probing exchanges with a smaller number of pupils in a given lesson, ensuring that over a longer period - say a day or a week rather than a single lesson - every pupil had such an opportunity. As a matter of fact, they too believed in maximising student participation, but the maximisation they aimed for was qualitative rather than quantitative.

Such extended and reciprocal rather than brief and one-sided exchanges, in which answers are justified, probed and built on, are more likely to achieve the goal of scaffolded understanding than round-the-class question and answer, but they require commitment to a second principle, that of collectivity, the idea that students are not lonely individuals in a crowd as famously portrayed from American elementary school classrooms by Philip Jackson (1967), but members of a learning community who listen to and learn from each other. They speak as co-contributors to a shared enterprise, not as individuals whose utterances are tacitly or explicitly in competition. This approach also runs strongly counter to the Anglo-American belief in individualised and differentiated learning, in which teacher-pupil interaction is essentially a private conversation within a whole-class context, but it sits well with Vygotsky, social constructivism and meta-cognition.

I have two video clips which I sometimes use to illustrate the success of this approach to applying lessons of international comparative analysis, one from Kursk in southern Russia and the other from England's rural North Yorkshire. On the face of it, the teaching is utterly different. The Kursk teacher retains the brisk, authoritative manner and clear social distance which run deep in Russian pedagogy, culture and history; the English teacher operates in a more informal, approachable and - yes - more English way. But what they have in common is the way they manage time and events to create space for teacher-student exchanges which are extended and carefully structured and to which the rest of the class have learned that they must listen intently, and it must be emphasised that the North Yorkshire teacher started moving in this direction only after she had watched, pondered and discussed the Kursk tape and transcript and others from the same dataset. For her, the idea of staying with one child over several cumulative exchanges ran sharply counter to her inherited belief, and to the wider 'folk pedagogy' (Bruner 1996) of English primary education, that in a given lesson every child must say something. Yet despite having made this fundamental change, her teaching remained emphatically English rather than Russian. She - and the children, for they too needed to be re-socialised pedagogically - applied and domesticated the principles but did not copy the practice. For not only was the practice visibly alien, but transplanted as a whole it would have conveyed values that the English teacher would not have wished to emulate (for example about the nature of knowledge and teacher authority) along with those that she was keen to apply.

Current efforts to reform classroom talk - which, after all is the heart of pedagogy - provide, I think, a good example of principled policy learning from international comparison. Central to our approach are not specific strategies to be copied, although some specific classroom indicators are proposed for teachers to consider as options, but instead five fundamental pedagogical values which together provide the core criteria for judging whether classroom talk, however it is organised and orchestrated, is genuinely dialogic: interaction must be *collective, reciprocal, supportive, cumulative* and *purposeful*. These now inform the

work of many UK teachers, especially in those local authorities like Barking and Dagenham, Bolton and North Yorkshire where we've been running dialogic teaching R & D projects (Alexander 2005a, 2005b), and an independent evaluation led by Adam Lefstein of Ben Gurion University confirms that in the East London schools there's now a signature pedagogy which is utterly distinctive in terms of the character and impact of teachers' and students' classroom talk (Lefstein and Snell 2011).

All of the teachers involved in these projects have striven to enact what is arguably the bottom line for educationally productive classroom talk, as captured by American academic Martin Nystrand and the great Russian philosopher Mikhail Bakhtin:

> What ultimately counts is the extent to which teaching requires students to think, not just to report or repeat someone else's thinking. (Nystrand *et al* 1997, 72)

> If an answer does not give rise to a new question from itself, it falls out of the dialogue. (Bakhtin 1986, 68).

Thus we square the circle: uncover and apply the *principles* of successful teaching elsewhere, don't merely copy the practice; test and validate these against the latest international evidence on how children most productively learn and teachers most effectively teach; and demonstrate that the result both increases student engagement and raises standards as measured by official tests.

And the circle can indeed be squared, which in these test-obsessed times is a prime political imperative. Sponsored by the American Educational Research Association (AERA), a conference at Pittsburgh University in September 2011 brought together leading international researchers on children's linguistic and cognitive development and on pedagogy and classroom talk. On the basis of the papers presented and reviewed, the conference judged that we now possess a critical mass of robust evidence demonstrating that classroom talk that is well-structured, reciprocal and cognitively challenging, and which immerses the student in a subject's distinctive conceptual and linguistic architecture and modes of enquiry, reasoning and argument, has a clear and positive impact on student attainment in English, mathematics and science as assessed by conventional tests. (Resnick, Asterhan, Clarke and Hofkens 2012). Further, and critical in relation to this paper's concerns, although schools and classrooms vary considerably, the principles of educationally productive talk appear to travel without too much difficulty from one classroom and culture to another.

Yet while the first and politically most influential of the three McKinsey reports concedes that good teaching matters, it announces:

> We have chosen *not* to focus on pedagogy or curricula, however important these subjects might be in themselves. These subjects are well-debated in the literature. (Barber and Mourshed 2007, 8)

If studies from the Type II stable stand or fall on the validity and reliability of the student attainment measures by which they define the performance of schools, education systems or jurisdictions, then in seeking to understand what *makes* schools or systems 'high performing', such studies also stand or fall on their capacity to engage in a conceptually valid and empirically defensible way with what schools and teachers do with the students whose attainment they seek to advance. It simply isn't good enough, in a study entitled *How the world's best performing systems come out on top* - where the word 'how' surely signals the intention to *explain* - to say, 'The quality of teaching is what makes the most difference, but we are not going to discuss teaching or define quality.' What kind of an explanation is that?

CONCLUSION

The example of the talk that lies at the heart of pedagogy takes us towards an answer to my initial question about why the ideas of Comenius – to take one instance - travelled so far and so successfully from their roots in 17th Century Moravia and are still discernible in the pedagogies of many of the world's countries. It is true that Comenius produced prescriptive textbooks to advance his theories of education and pedagogy, but it is the *principles* that proved more durable. And how revolutionary they were at a time when education was largely confined to boys from wealthy families and Latin was the medium of instruction. Thus, on the nature of the education system:

- There should be universally-available education for both boys and girls, and for all children regardless of income or social position.
- The school system should start with education within the family up to the age of six, and be followed by elementary school, secondary school and university.
- Teaching should be in the pupils' mother tongue.
- Lifelong education should be available for adults. (Pánková 2010)

And on pedagogy (from Keatinge's 1896 translation of Comenius's *Great Didactic*):

> There should be one teacher for each class ... Time should be carefully divided, so that each day and each hour may have its appointed task ... The same exercise should be given to the whole class ... Everything should be taught thoroughly, briefly, and pithily, that the understanding may be, as it were, unlocked with one key ... All things that are naturally connected ought to be taught in combination ... Every subject should be taught in definitely graded steps, that the work of one day may thus expand that of the previous day, and lead up to that of the morrow. (Comenius 1896 [1657], 312-334)

And foreshadowing the dialogic principles referred to above:

> The scholars should be given leave to ask questions on any point that they wish explained ... [They] should ask the question openly. In this way the whole class will benefit, as much by the question as by the answer.' (Comenius 1896 [1657], 320).

There, in a treatise from a time and place far removed from ours, are two of the principles of dialogic teaching: collectivity and reciprocity. If we add to such principles the illustrated textbooks and the use of natural and man-made objects to stimulate children's interest and advance their understanding – objects we can still see in the Comenius pedagogical museum in Prague and the A.H.Francke museum in Halle in Saxony – we begin to understand the continuing significance of this founder of modern pedagogy.

Mention of Germany and the Czech Republic, neither of which country has a seat at the PISA top table or is deemed by the current wave of policy borrowers to have educational systems, policies or practices from which we can learn, takes me to my closing questions. Others must answer them.

First, we know that Singapore, Hong Kong, Korea, Shanghai-China and Finland do well in PISA tests of student attainment in reading, maths and science. But what else do their students learn, and how well? Do these systems provide their children with an education which is about significantly more than passing tests in three subjects? And if the wider curriculum in top-performing PISA systems were to be measured as assiduously as the reading, maths and science are measured, would the same countries still head the league table?

Second, we are being asked to subscribe to a view of 'world class' schooling as *outperforming* other countries. The uncomplicated but undeniably imperialist goal, shared by the policy supremacists and their unacknowledged alter ego Nigel Molesworth, is *How to be Topp.* But in a world facing the crises of climate change, resource depletion, over-population,

environmental degradation and geo-political instability, is this really how 'world class' should be defined? Should we not consider the merits of 'world class' as *sustaining* the world rather than beating it? As fostering international interdependence rather than national supremacy? Whatever happened to education's *moral* purposes?

Third, if we take this rather different view of education, adopting a broader, richer and more humane vision than that espoused by McKinsey's executives and their Westminster and Washington cheerleaders, have we selected the right countries from which to learn about what makes for a 'world class' education?

Fourth, is it not at least possible that this whole current PISA-led obsession does a grave injustice to countries, schools and teachers that care no less passionately about educational standards but for whom standards mean more than test performance - and ultimately therefore to the cause of education itself?

Finally, national curricula are increasingly being reconfigured to respond less to national culture, values and needs than the claims of 'international benchmarking' and 'world class' educational standards – the latter equated with test scores in a limited spectrum of human learning. The curriculum narrows to what is tested, the summative function of assessment is elevated over its formative contribution to children's understanding and progress, and the larger questions of purpose and value, which in democratic societies ought to be central to educational debate, are neglected. With PISA 2012 now in progress, is this not a good time to redress the balance?

REFERENCES

Alexander, R.J. (1996) *Other Primary Schools and Ours: hazards of international comparison,* Warwick: University of Warwick, CREPE.

Alexander, R.J. (1997) *Policy and Practice in Primary Education: local initiative, national agenda,* London: Routledge.

Alexander, R.J. (2001) *Culture and Pedagogy: international comparisons in primary education*, Oxford: Blackwell.

Alexander, R.J. (2005a) *Talk for Learning: the second* year, Northallerton: North Yorkshire County Council.

Alexander, R.J. (2005b) *Teaching Through Dialogue: the first year*, London: London Borough of Barking and Dagenham.

Alexander, R.J. (2006) *Education as Dialogue: moral and pedagogical choices for a runaway world*, Hong Kong: Hong Kong Institute of Education with Dialogos.

Alexander, R.J. (2008a) *Essays on Pedagogy*, Abingdon: Routledge.

Alexander, R.J. (2008b) *Education for All, the Quality Imperative and the Problem of Pedagogy* (CREATE Research Monograph 20), Universities of Sussex and London, with the Department for International Development: Consortium for Research on Educational Access, Transitions and Equity. http://www.create-rpc.org/pdf documents/PTA20.pdf

Alexander, R.J. (2008c) *Towards Dialogic Teaching: rethinking classroom talk*, York: Dialogos.

Alexander, R.J. (2009) 'Towards a comparative pedagogy', in Cowen, R. and Kasamias, A.M. (ed) *International Handbook of Comparative Education,* New York: Springer, 911-929.

Alexander, R.J. (2010a) ' "World class schools" – noble aspiration or globalised hokum?', BAICE Presidential Address at the 10th UKFIET Conference on Education and Development, *Compare: a journal of comparative education*, 40(6), 801-817.

Alexander, R.J. (ed) (2010b) *Children, their World, their Education: final report and recommendations of the Cambridge Primary Review*, Abingdon: Routledge.

Alexander, R.J. (2011) 'Legacies, policies and prospects: one year on from the Cambridge Primary Review', *Forum* 53(1), 71-92.

Alexander, R.J., Doddington, C., Gray, J., Kershner, R. (ed) (2010) *The Cambridge Primary Review Research Surveys*, Abingdon, Routledge.

Bakhtin, M.M. (1986) *Speech Genres and Other Late Essays*, Austin: University of Texas Press.

Bangs, J. (2011) Oral statement to the New Visions in Education Group, London.

Barber, M. and Mourshed, M. (2007) *How the World's Best-Performing School Systems Come Out on Top*, Dubai: McKinsey and Company.

Barber, M., Whelan, F., Clark, M. (2011) *Capturing the Leadership Premium: how the world's top school systems are building leadership capacity for the future*, Dubai: McKinsey and Company.

Benavot, A. (2008) 'The organisation of school knowledge: official curricula in global perspective', in Resnick, J. (ed) *The Production of Knowledge in the Global Era*, Rotterdam and Taipei: Sense Publishers.

Benavot, A., Cha, Y-K., Kames, D., Meyer, J.W., Wong, S-Y (1991), 'Knowledge for the masses: world models and national curricula, 1920-1986', *American Sociological Review,* 56.
Bernstein, B. (1970) 'Education cannot compensate for society', *New Society*, 26 February, 344-7.
Boyer, E.L. (1983) *High School: a report on secondary education in America*, New York: Harper & Row, p. 6, quoted in H.J. Noah 'The use and abuse of comparative education', in Altbach, P.G., Kelly, G.P. (ed) *New Approaches to Comparative Education,* Chicago: University of Chicago Press.
Bruner, J.S. (1996) *The Culture of Education*, Cambridge MA: Harvard University Press.
Coffield, F. (2012) 'Why the McKinsey reports will not improve school systems', *Journal of Education Policy*, 27(1), 131-149.
Comenius, J.A. [1657] translated by M.W.Keatinge (1896), *The Great Didactic of Comenius,* London: A. and C. Black.
Department for Education (2010a) *Remit for the Review of the National Curriculum in England*, London: DfE.
Department for Education (2010b) *The Importance of Teaching: the schools White Paper 2010*, London: DfE.
Department for Education (2011a) *The Framework for the National Curriculum: a report by the Expert Panel for the National Curriculum Review*, London: DfE.
Department for Education (2011b) *Review of the National Curriculum in England: what can we learn from the English, mathematics and science curricula of high-performing jurisdictions*? London: DfE.
Gage, N. (1978) *The Scientific Basis of the Art of Teaching*, New York: Teacher's College Press.
Galton, M. and Simon, B. (ed) (1980) *Progress and Performance in the Primary Classroom*, London: Routledge.
Gray, J., Galton, M., McLaughlin, C., Clarke, B. and Symonds, J. (2011) *The Supportive School: wellbeing and the young adolescent*, Newcastle upon Tyne: Cambridge Scholars Publishing.
Hargreaves, D.H. (2011) *A Self-Improving School System in International Context*, Nottingham: National College for School Leadership.
Hogan, D. with Towndrow, P., Rahim, R.A., Chan, M., Luo, S., Zher. S.Y.T., Asye, K.M., Kiang, T.T., Chen, L.S., Kwek, D. amd Kaur, B. (2012, forthcoming) *Core 1&2 Interim Report on Instructional Practices in Singapore in Secondary 3 Mathematics and English, 2004 and 2010*, Singapore: National Institute of Education.
Jackson, P.W. (1967) *Life in Classrooms*, New York: Holt, Rinehart and Winston.
Lefstein, A. and Snell, J. (2011) 'Classroom discourse: the promise and complexity of dialogic practice' in S.Ellis, E.McCartney and J.Bourne (eds) *Insight and Impact: applied linguistics and the primary school*, Cambridge: Cambridge University Press.
Li, J. (2012) *Cultural Foundations of Learning: east and west*, New York: Cambridge University Press.
Mortimore, P., Sammons, P., Stoll, L., Lewis, D., Ecob, R. (1988) *School Matters: the Junior Years,* London: Open Books.
Mourshed, M., Chijioke, C., Barber, M. (2010) *How the World's Most Improved School Systems Keep Getting Better*, Dubai: McKinsey and Company.
National Research Council (2003) *Understanding Others, Educating Ourselves: getting more from international comparative studies in education* (C.Chabott and E.J.Eliott, editors), Washington DC: The National Academies Press, http://www.nap.edu/openbook.php?isbn=0309088550
Northen, S. (2011) 'Follow my (associate) leader', *The Guardian*, 15 November.
Nystrand, M., Gamoran, A., Kachur, R., Prendergast, C. (1997) *Opening Dialogue: understanding the dynamics of language and learning in the English classroom*, New York: Teachers College Press.
Oates, T. (2010) *Could Do Better: using international comparisons to refine the national curriculum in England*, Cambridge Assessment.
OECD (2010a) *PISA 2009 Results: learning trends*, Paris: OECD.
OECD (2010b) *PISA 2009 Results: what makes a school successful?* Paris: OECD.
OECD (2010c) *PISA 2009 Results: overcoming social background,* Paris: OECD
OECD (2010d) *PISA 2009 Results: executive summary*, Paris: OECD.
OECD (2011a) *PISA 2009 Results: quality time for students – learning in and out of school*, Paris OECD.
OECD (2011b) *PISA 2009 Results: against the odds*, Paris: OECD.
OECD (2012) 'What PISA assesses', http://www.pisa.oecd.org/pages/0,3417,en_32252351_32235918_1_1_1_1_1,00.html (accessed February 2012).
Osborn, M., Broadfoot, P., McNess, E., Planel, C., Ravn B. and Triggs, P. (2003) *A World of Difference? Comparing learners across Europe*, Maidenhead: Open University Press.
Pánková, M. (ed) (2010) *The Legacy of J.A.Comenius: traditions and challenges of the Czech culture and education to Europe*, Prague: Pedagogical Museum of J.A.Comenius.
Raffe, D. (2011a) 'What can we learn from other school systems about improvement? Comments on *How the world's most improved school systems keep getting better*, presentation to Parliamentary Research Enquiry, 28 June.
Raffe, D. (2011b) 'Policy borrowing or policy learning? How (not) to improve education systems', *CES Briefing,* 57.
Resnick, L.B., Asterhan, C., Clarke, C. and Hofkens, T. (eds) (2012, forthcoming)*Socializing Intelligence* [papers from the September 2011 AERA conference 'Socializing intelligence through academic talk and dialogue' held at the University of Pittsburgh], Washington DC: AERA.
Reynolds, D. and Farrell, S. (1996) Worlds Apart? A review of international surveys of educational achievement involving England, London: Ofsted.

Sadler, M. [1900] 'How can we learn anything of practical value from the study of foreign systems of education?' In J.H.Higginson (ed) *Selections from Michael Sadler: studies in world citizenship.* Liverpool: Dejall and Meyorre.
Tobin, J.J., Hsueh, Y., Karasawa, M. (2009) *Preschool in Three Cultures Revisited*, Chicago: University of Chicago Press.
Tobin, J.J., Wu, D.Y.H., Davidson, D.H. (1989) *Preschool in Three Cultures: Japan, China and the United States*, New Haven: Yale University Press.
UNICEF (2007) *Child Poverty in Perspective: an overview of child well-being in rich countries*, Innocenti Report Card 7, Florence: UNICEF Innocenti Research Centre.
UNICEF (2011) *Child Well-being in the UK, Spain and Sweden: the role of inequality and materialism*, London: UNICEF UK.
Watkins, D.A. and Biggs, J.B. (1996)*The Chinese Learner: cultural, psychological and contextual differences*, Hong Kong: Comparative Education Research Centre.
Wilkinson, R. and Pickett, K. (2010). *The Spirit Level: why equality is better for everyone,* London: Penguin.
Wolf, A. (2002) *Does Education Matter? Myths about education and economic growth*, London: Penguin Books.